

1636

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/747,385

DATE: 05/10/2001
TIME: 08:37:52

Input Set : A:\Uc-998-1.app
Output Set: N:\CRF3\05102001\I747385.raw

#4
KFD
S-a3-01

ENTERED

3 <110> APPLICANT: Attarian, Gwynne
4 Podkaminer, Kara K.
5 Yoder, Sean C.
6 Kinder Haake, Susan A.
7 The Regents of the University of California
9 <120> TITLE OF INVENTION: Fusobacterium Nucleic Acids, Plasmids and Vectors
11 <130> FILE REFERENCE: 02307E-099810US
13 <140> CURRENT APPLICATION NUMBER: US 09/747,385
14 <141> CURRENT FILING DATE: 2000-12-22
16 <150> PRIOR APPLICATION NUMBER: US 60/173,168
17 <151> PRIOR FILING DATE: 1999-12-27
19 <160> NUMBER OF SEQ ID NOS: 21
21 <170> SOFTWARE: PatentIn Ver. 2.1
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 407
25 <212> TYPE: PRT
26 <213> ORGANISM: Fusobacterium nucleatum
28 <220> FEATURE:
29 <223> OTHER INFORMATION: RepA
31 <400> SEQUENCE: 1
32 Met Asp Phe Ser Ser Ile Lys Lys Ser Leu Gly Leu Ile Asn Phe Arg
33 1 5 10 15
35 Asp Leu Lys Lys Tyr Ile Leu Gly Leu His Gln Lys Leu Gly Asn Leu
36 20 25 30
38 His Ile Thr Asn Ile Thr Asn Lys Lys Ile Glu Thr Ile Phe Leu Phe
39 35 40 45
41 Glu Lys Phe Ile Asn Asp Leu Asp Asn Asn Thr Leu Thr Ile Arg Val
42 50 55 60
44 Thr Lys Asp Ser Leu Tyr Phe Phe Asn Ile Ala Asn Ser Tyr Leu Arg
45 65 70 75 80
47 Phe Leu Phe Ser Asp Val Arg Lys Leu Ser Gly Lys Tyr Ser Lys Leu
48 85 90 95
50 Leu Val Pro Tyr Leu Met Glu Phe Ser His Lys Lys Glu Ala Glu Phe
51 100 105 110
53 Glu Lys Glu Arg Phe Phe Asn Ile Leu Glu Val Glu Glu Ser Tyr Arg
54 115 120 125
56 Asn Asn Leu Ser Asp Phe Asn Lys Arg Ile Leu Lys Pro Ala Val Glu
57 130 135 140
59 Glu Leu Lys Thr Leu Phe Glu Asn Leu Lys Val Glu Arg Leu Lys Asn
60 145 150 155 160
62 Gly Arg Val Ile Lys Gly Tyr Lys Phe Ser Trp Thr Asn Asp Phe Asn
63 165 170 175
65 Phe Gln Asn Lys Lys Asp Asn Ile Glu Glu Ala Glu Val Val Glu Glu
66 180 185 190
68 Lys Glu Asn Ile Ala Ser Gly Glu Leu Glu Lys Tyr Phe Lys Ser Thr
69 195 200 205
71 Phe Thr Asp Val Asn Tyr Ser Lys Lys His Lys Glu Val Leu Glu Lys

RAW SEQUENCE LISTING

DATE: 05/10/2001

PATENT APPLICATION: US/09/747,385

TIME: 08:37:52

Input Set : A:\Uc-998-1.app

Output Set: N:\CRF3\05102001\I747385.raw

```

72      210      215      220
74 Leu Leu Lys Asn Asn Ser Leu Glu Tyr Ile Lys Lys Tyr Leu Ser Glu
75 225      230      235      240
77 Gln Trp Glu Tyr Val Gln Asn Asp Lys Asn Ile Leu Asn Lys Ser Ala
78      245      250      255
80 Tyr Phe Ser Lys Leu Ile Leu Glu Glu Lys Ala Val Tyr Lys Asn His
81      260      265      270
83 Leu Pro Ala Asp Tyr Glu Glu Leu Lys Val Glu Glu Arg Asn Arg Asn
84      275      280      285
86 Ile Glu Ser Thr Asn Thr Ile Thr Ser Leu Lys Asp Leu Val Glu Lys
87      290      295      300
89 Asp Ile Thr Asp Tyr Glu Val Arg Lys Asn Ile Thr Pro Glu Gln Ile
90 305      310      315      320
92 Glu Gln Glu Val Leu Phe Lys Ile Asp Val Thr Glu Glu Glu Tyr Asn
93      325      330      335
95 Lys Ile Lys Glu Asp Trp Ile Ile Lys Arg Lys Asp Glu Val Pro Asn
96      340      345      350
98 Ser Asp Pro Lys Leu Leu Glu Ile Ile Phe Asn Ala Ser Gln Ser Lys
99      355      360      365
101 Lys Tyr Asn Ile Ile Asn Thr Lys Glu Glu Val Asn Glu Lys Glu Lys
102      370      375      380
104 Glu Leu His Glu Leu Glu Asn Ile Lys Arg Met Gln Glu Glu Leu
105 385      390      395      400
107 Asn Lys Leu Lys Lys Glu Val
108      405
111 <210> SEQ ID NO: 2
112 <211> LENGTH: 1224
113 <212> TYPE: DNA
114 <213> ORGANISM: Fusobacterium nucleatum
116 <220> FEATURE:
117 <223> OTHER INFORMATION: repA
119 <400> SEQUENCE: 2
120 atggattttt cttctataaa aaaaagttaa ggtttaatta attttagaga tttaaaaaaa 60
121 tatatttttag gacttcacata aaaattagga aatttacata ttactaatat aacaaataaa 120
122 aaaattgaaa caatcttttt atttgaaaaa ttcataaatg atttagataa taatacttta 180
123 actataagag taacaaaaga ttctctttat tttttaataa ttgctaacag ttatttaagg 240
124 tttctctttt cagatgtagt aaaactttca ggaaaatatt caaagttatt gggttccttat 300
125 ttaatggagt ttagtcataa aaaagaagct gaatttgaaa aagagagatt ttttaattatt 360
126 ctagaagttg aagaaaagta tagaaataat ttatcagatt ttaataagag aattctaaaa 420
127 ccagctgttg aagaattaaa aacacttttt gaaaatttaa aggttgagcg attaaaaaat 480
128 ggaagagtaa taaaaggata taaatttagc tggactaatg attttaattt tcaaaataag 540
129 aaagataata tagaagaagc agaagtagtg gaagaaaaag aaaatattgc ttcaggagag 600
130 ttagaaaaat attttaaatc aacttttact gatgtaaatt attcaaagaa gcataaagaa 660
131 gtttttagaaa aattattaaa aaataatagt ttagaatata ttaaaaaata tttatctgag 720
132 cagtgggagt atgtacaaaa tgataaaaat attttaaata aatcagcata tttctcaaaa 780
133 ctaatttttag aagaaaaagc agtatataaa aatcatctac cagctgacta tgaagaacta 840
134 aaagttgaag aaagaaatag aaatatagaa agtacaaata ctattacatc attaaaagat 900
135 ttagtagaaa aagacattac agattatgaa gttagaaaga atataactcc tgaacaaata 960
136 gaacaagaag ttttatttaa aatagatgta actgaagaag aatataataa gattaaagaa 1020

```

RAW SEQUENCE LISTING

DATE: 05/10/2001

PATENT APPLICATION: US/09/747,385

TIME: 08:37:52

Input Set : A:\Uc-998-1.app

Output Set: N:\CRF3\05102001\I747385.raw

```

137 gattggataa taaaacgaaa agatgaagtt cctaatagtg atccaaaact tttagaaatt 1080
138 atattttaatg caagtcaatc aaaaaaatat aatataatta atactaaaga agaagttaat 1140
139 gaaaaagaaa aagagcttca cgaattagaa gaaaatataa aaagaatgca agaagaacta 1200
140 aataaattaa aaaaagaggt atag 1224
143 <210> SEQ ID NO: 3
144 <211> LENGTH: 22
145 <212> TYPE: DNA
146 <213> ORGANISM: Artificial Sequence
148 <220> FEATURE:
149 <223> OTHER INFORMATION: Description of Artificial Sequence: iteron sequence
150 within the origin of replication of plasmid pFN1
152 <400> SEQUENCE: 3
153 tcaactttaa caggacaaat tt 22
156 <210> SEQ ID NO: 4
157 <211> LENGTH: 132
158 <212> TYPE: DNA
159 <213> ORGANISM: Artificial Sequence
161 <220> FEATURE:
162 <223> OTHER INFORMATION: Description of Artificial Sequence: six copies of
163 the iteron sequence within the origin of
164 replication of plasmid pFN1
166 <400> SEQUENCE: 4
167 tcaactttaa caggacaaat tttcaacttt aacaggacaa attttcaact ttaacaggac 60
168 aaattttcaa ctttaacagg acaaattttc aactttaaca ggacaaattt tcaactttaa 120
169 caggacaaat tt 132
172 <210> SEQ ID NO: 5
173 <211> LENGTH: 1230
174 <212> TYPE: DNA
175 <213> ORGANISM: Artificial Sequence
177 <220> FEATURE:
178 <223> OTHER INFORMATION: Description of Artificial Sequence: repA homolog
179 sequence of plasmid pAD52
181 <400> SEQUENCE: 5
182 atggattttt cttctataaa aaaaagttta ggtttaatta attttagaga tttaaaaaaa 60
183 tatatttttag gacttcatca aaaattagga aatttacata ttactaatat aacaaataaa 120
184 aaaattgaaa caatcttttt atttgaaaaa ttcataaatg atttagataa taatacttta 180
185 actataagag taacaaaaga ttctctttat ttttttaata ttgctaacag ttatttaagg 240
186 tttctctttt cagatgtagt aaaactttca ggaaaatatt caaagttatt gggttccttat 300
187 ttaatggagt ttagtcataa aaaagaagct gaatttgaaa aagagagatt ttttaatat 360
188 ctagaagttg aagaaagtta tagaataaat ttatcagatt ttaataagag aattctaaaa 420
189 ccagctgttg aagaattaaa aacacttttt gaaaatttaa aggttgagcg attaaaaaat 480
190 ggaagagtaa taaaaggata taaatttagc tggactaatg attttaattt tcaaaataag 540
191 aaagataata tagaagaagc agaagtagtg gaagaaaaag aaaataaaaa tattgctcct 600
192 ggagagttag aaaaatatat taaaacaact ttccctggtg taaattattc aaagaagcat 660
193 aaagaagttt tagaaaaatt attaaaaaat aatagtttag aatatattaa aaaatattta 720
194 tctgagcagt gggagtatgt acaaaacgat aaaaatatat taaataaatc agcatatttt 780
195 tcaaaactaa tcttagaaga aaaagcagta tataaaaatc atctaccagc tgactatgaa 840
196 gaattaaaag ttgaagaaag aaatagaaat atagaaagta caaatactat tacatcatta 900
197 aaagatttag tagaaaaaga cattacagat tatgaattta gaaagaatat aactcctgaa 960

```

RAW SEQUENCE LISTING

DATE: 05/10/2001

PATENT APPLICATION: US/09/747,385

TIME: 08:37:52

Input Set : A:\Uc-998-1.app

Output Set: N:\CRF3\05102001\I747385.raw

```

198 caaatagaac aagaagtttt attttaaata gatgtaactg aagaagaata taataagatt 1020
199 aaagaagatt ggataataaa acaaaaagaa gtagttccta atagtgatcc agaactttta 1080
200 gaagttatat ttaatgcaag tcaatcaaaa aaatataata taattaatac taaagaagaa 1140
201 gttaatgaaa aagaaaaaga gcttcacgaa ttagaagaaa atataaaaag aatgcaagaa 1200
202 gaaataaata aattaaaaaa agaggtatag 1230
205 <210> SEQ ID NO: 6
206 <211> LENGTH: 5887
207 <212> TYPE: DNA
208 <213> ORGANISM: Artificial Sequence
210 <220> FEATURE:
211 <223> OTHER INFORMATION: Description of Artificial Sequence: plasmid pFN1
213 <400> SEQUENCE: 6
214 catataataa ctttttgttt ctttcttcta gtactttttt ttctaactct ttttaatttct 60
215 tttttactttt ttcaattttt tctaattcaa tcgcttttaa ttctttctaaa gtcttttttaa 120
216 ttttttttagt agctttccata catatcacac tccagcatta ttatttataa aaatataatt 180
217 atataataca tatctagttaa aataaatcaa gtagtgcggg cttaaacaag agccatataa 240
218 attaaaaaat atatataattt ttcttttaat ttctaataata aaatgggtata atattttata 300
219 tggcattgtt taaggcaagc tactcactcc tgccgctggt tagcttgctg ggaagtatcc 360
220 cagaccctca gaagcaccta caaaaattaa aaatatattt ttaattttct tcggtactta 420
221 taggggggatg acccctataa cccctgcgag ttgattttat aaaattcttg cagggacagc 480
222 tgcaattatt ttataaaacc aactttaaag agagtataa taatgattaa atttacatta 540
223 agattaataa taatgattaa atttacatta agattaacgg aagatgaaaa aaaactttta 600
224 gatataaaa ctgatgaatt aggtaaatca aaaaatgaag ttttaaagtt tcttataaac 660
225 aataaattgg aagatactaa aaaagaattt gacctattaa atgagcttga taaaaattat 720
226 aaagagctag gttttcagat taaaaaatt ggagtagttt taaatcagat taataaaaaat 780
227 ttttatgaag ataagaaaat acagattgaa gaaatccaag gagcggtaga tgaattatgg 840
228 cagtctataa aagtgtcaaa ggagtaggaa aaactaaaag cagtttatac aatattttta 900
229 agtatgtagg aagtcaaaat gaaaaagaaa aagatgatag agtttataaa actactggta 960
230 taaatgttag tgatgattat aaaaaagcct ttaaagaaat gatgttaaca aaagagcttc 1020
231 attgtaagtt agacggtaga caatatagac accatattca atcttttaaa cctgggtgaag 1080
232 tagatgaaga aacagcacat aaaaatggcag tagaatttgc agaaaaaaat tttaaaggct 1140
233 ttgatgtttt tatatcaact cacattgata aaggacacat acataacat ataattatta 1200
234 atactgttaa tattgatact ggaatgaagt tcagagaatt aaataagaat gaatataatc 1260
235 aaaaaaaga aaataatgtt gaattaaaat ctacgaatt ttatttagaa gatttaaaaa 1320
236 aatctagtga tgaaatttgt cttgcaaata atttatcggg gatcccccaa aaagaaaaag 1380
237 ctgaaaagtc gaattattat aatagacgag aatataatgt tgtgatgaac aaaacaagtt 1440
238 ataaaatgga gctggcaaaa gatataaaaa gagcttccaa gaattgtaaa tcaaaagaag 1500
239 attttataaa agcattagat gaaaaagggt ttattgtgga ttgggaagac cataaaaaac 1560
240 atataacttt taaatttaaa gatgagaaaa agaaatcaat tagattagca aatttagaaa 1620
241 aaactttcca agatgaaact tttaaaaagg aatacctgga acagcaattt ttaataaatc 1680
242 aaaaaactga agaaatagga aatttcaaaa ttaaagttaa tactgaagct gaaaaaacta 1740
243 atgaagaaaa atatcaagag cttctaaata aaaagagaga acaagataaa ttaatagctg 1800
244 aagaaaaatt gaaaaagcaa aaagaagaaa ttgagaaaaa gaaaaattta aatagaaaac 1860
245 aagggtttcgg aataggagat taaaaatggc tatatttagat gatgatgtaa atgaagttaa 1920
246 aaatgtgaat gagcaagaaa taaaaatgga tagttatatt aaaaaataa ttgaaaatgt 1980
247 tttagaagtt cagctaaaag aacataaaga aatagcttcc attgctaaaa ctaaaatagc 2040
248 tgaagtaact ttagaactag aaaaattaaa acagctggag aaagcaacta ctaaaatag 2100
249 agatgataca aatataatta caaataaaat gattgaaaat gttgaaaatt ataataaagt 2160
250 atttttagaa agaattgata aatttaattt attgatggta gaaaagttaa atgaagtaaa 2220

```

RAW SEQUENCE LISTING

DATE: 05/10/2001

PATENT APPLICATION: US/09/747,385

TIME: 08:37:52

Input Set : A:\Uc-998-1.app

Output Set: N:\CRF3\05102001\I747385.raw

```

251 aaataactaat caaatatttg ctgagacatt agataaatca gatattgttg ataacccttaa 2280
252 taagagtaga gcagaaatat ttgagaaatt ttcaactgat gttgctacaa aaacaaataa 2340
253 tacttttaaat ttaattgaag atagtctaaa taaattaaaa atagccttct atacagctgt 2400
254 atctgttatt gtaatatattt tatttttttac ggggataatt ttatataaga caaataatag 2460
255 agttgctagt gttgaagaaa gcttaaataa tatatctagt tcagtaactg gattagttaa 2520
256 aggggactta aagttttggt acagtgaaga agacaaaaaa gcctatgtaa gtaatgtaga 2580
257 aagtattaaa aaaaaataag atagcaagaa gaaaaataa aagcttcagt aagataaaaa 2640
258 agcaataaac atttaattta ttgctttttt attttatagt ttagtcattg agggtaaatt 2700
259 tttatagtaa tatatatattt aacaatttta ctatattact ttttaacatt ctttagaaac 2760
260 atatccataa tatagttcat tagacttgcg acagttatct catttgtagc agcatacttt 2820
261 ttgaaatttg agtaaattctc tgagtttggt ttcatgata tagttatattc atttttttaa 2880
262 gttctataat gttcaggaa gacacagtt tcattattta cattaatttt tctttcgtta 2940
263 acaatattaa aaagtatttc taaattttct tcattaaaca aactattata tttttctgga 3000
264 actgtcaatt gaaattcttg cttttcatct atattacttt ttatattact atctatatta 3060
265 cttacacttt cagtaacttc ttcttcaggc ttacacaaat cttttaaagt ttcatacaaca 3120
266 taagtcataa cagctttttt tatccctttt ggagcataat taacttttac atcatctatg 3180
267 aatgtaacat cataagttcc attctctagt ttgtaatct cattaaatttt ttctaaatat 3240
268 tttttagcaa gacccaaacc ttgtaccata aaaaatcact ccccttttta tcttagtaat 3300
269 atatattact ataattattc taaattttta catagcataa atacacaagt tcaatcggta 3360
270 atattttatta ctacaattat actttattac ttcaacagta tacacctgtt tttataaaaa 3420
271 ataaagaggt ttttttatgc agaaaaataa ttaaaaataa tttataaaaa aatctctaaa 3480
272 atacttttta aggtttttt ttttattgag ccatactttt tattgttaaa atgtctaaaa 3540
273 tcatttttag gggatcccta ggacttttaa attgatttta aatgctgttt gtgtgttaaa 3600
274 cttctttatg tttttttaa taaaaaaagt taggcattgt gagagtccta actttttatg 3660
275 tcgttttggt caagcaacgg atactttgtt gctatgtttc aaactaatta tatcacattc 3720
276 aatttttaaat ttcaatatgt tattaattct aattgtcgtt tcttccaaag gaggtgaaag 3780
277 atgttcaagg aatggatact attgataatg gctttaatat ttattgctat tttacaatat 3840
278 gcagttatat taatatgggt tcacattatc acaagtttat tagcatagtt cttgtaaaaa 3900
279 ataagtctag agagtttagt aagtttgaga gcttgaaaac tctctttttt tttagggaaa 3960
280 catattttata tttattttca atctttttta ccgaaagtgt attttccact ttcggtaatt 4020
281 ttttatattt tttatttgag cttcttttat taaaaaaat cacattacta atatgattaa 4080
282 tataataata ttatataata atatataata cagcactcat tttctttttt aatagcaatg 4140
283 taaaacaaaa agataacagg acaaattttc aactttaaca ggacaaattt tcaactttta 4200
284 caggacaaat tttcaacttt aacaggacaa attttcaact ttaacaggac aaattttcaa 4260
285 ctttaacagg acaaattttc aactttaaca ggacaaattt cattgacagt cttatattat 4320
286 tgggtgataa tgtttttatg aaataaaatt tccataaaag gagctgaaga ttttagtgaa 4380
287 taatgattta gtaaaagtac ataaagattt taccaaatta aatataggga cattaagtga 4440
288 aaaagaatta gaattatttt attatatatg tttaaatgta aaggatatta gagatgaaat 4500
289 tataacaatg gatttttctt ctataaaaaa aagtttaggt ttaattaatt ttagagattt 4560
290 aaaaaaatat attttaggac ttcatcaaaa attaggaaat ttacatatta ctaatataac 4620
291 aaataaaaaa attgaaacaa tctttttatt tgaaaaattc ataaatgatt tagataataa 4680
292 tacttttaact ataagagtaa caaaagattc tctttatttt tttaattatt ctaacagtta 4740
293 tttaaggttt ctcttttcag atgttagaaa actttcagga aaatattcaa agttattggt 4800
294 tctttattta atggagttta gtcataaaaa agaagctgaa ttgaaaaag agagattttt 4860
295 taatattcta gaagttgaag aaagttatag aaataattta tcagatttta ataagagaat 4920
296 tctaaaacca gctgttgaag aattaaaaac actttttgaa aatttaaagg ttgagcgatt 4980
297 aaaaaatgga agagtaataa aaggatataa atttagctgg actaatgatt ttaattttca 5040
298 aaataagaaa gataatatag aagaagcaga agtagtgaa gaaaaagaaa atattgcttc 5100
299 aggagagtta gaaaaattt ttaaatcaac ttttactgat gtaaattatt caaagaagca 5160

```

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/747,385

DATE: 05/10/2001

TIME: 08:37:54

Input Set : A:\Uc-998-1.app

Output Set: N:\CRF3\05102001\I747385.raw

L:428 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14
L:429 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14
L:430 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14
L:431 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14
L:432 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14
L:433 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14
L:434 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14
L:435 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14
L:436 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14